A multimedia interactive conferencing application for personal workstations

- Robinson, J.; Rubinov, E.; Toulson, C.; Prasada, B.; Sabri, S.; Goldberg, N.; Vonderweidt, G.
Dept. of Syst. Design Eng., Waterloo Univ., Ont., Canada

This Paper Appears in :
Communications, IEEE Transactions on

on Pages: 1698 - 1708

Nov. 1991 Vol. 39 Issue: 11 ISSN: 0090-6778
References Cited: 31
CODEN: IECMBT
Accession Number: 4108823

Abstract:

The multimedia interactive conferencing application (MICA), a personal-workstation application for multipoint visual teleconferencing, is described. MICA allows people at two or more locations to share visual material such as documents, photographs, and computer screens in a highly interactive way. It supports the distribution, storage, retrieval, and high-quality display of visuals, real-time interaction by pointing and annotation, and meeting services facilities. The context of multimedia teleconferencing and computer-supported cooperative work is established, relating earlier research to the design of MICA. The services MICA offers are outlined. The handling, compression, and display of multiple media, and the design of a suitable user interface for MICA are discussed.

Subject Terms:

A Multimedia Interactive Conferencing Application for Personal Workstations

John Robinson, *Member, IEEE*, Eliot Rubinov, Chris Toulson, Birendra Prasada, *Fellow, IEEE*, Shaker Sabri, *Member, IEEE*, Naftaly Goldberg, and Guy Vonderweidt

***Abstract*—The multimedia interactive conferencing application (MICA) is a personal-workstation application for multipoint visual teleconferencing. It allows people at two or more locations to share visual material such as documents, photographs, and computer screens, in a highly interactive way. It supports the distribution, storage, retrieval and high-quality display of visuals, real-time interaction by pointing and annotation, and meeting services facilities. In this paper we establish the context of multimedia teleconferencing and computer supported cooperative work, relating earlier research to our design of MICA. We outline the services MICA offers, and then focus on two of the major technical challenges: the handling, compression and display of multiple media, and the design of a suitable user interface. A third major area, the multipoint service that supports the application, is detailed in a companion paper.**

I. Introduction

A RECENT trend in multimedia image handling and display has been the move from dedicated systems with special-purpose hardware towards applications that run on conventional workstations. The graphic quality and processing power now available make it possible to handle text, graphics, scanned documents and natural (photographic) images within standard desktop personal computers. At the same time, research in multipoint communication has provided very powerful services for communicating between distributed workstations. These twin technology developments promise new solutions to a well identified application area—multimedia visual teleconferencing. Teleconferencing is not the only application that relies on advanced media-handling capabilities—others include publishing and multimedia databases—but it is particularly demanding in its use of dynamic interaction over a network.

The objective in teleconferencing is to provide a shared space wherein users can interact as they would in a face-to-face meeting. Visual teleconferencing is a subfield, defined as the sharing, in real time, of visual material by participants. It should be differentiated from the term videoconferencing which is reserved for full-motion video support between participating sites, allowing meeting participants to see each-other's nonverbal cues and giving a sense of presence to the meeting. The focus in visual teleconferencing, though, is on what participants talk about; for example, documents, slides and pictures. We have adopted the term visual teleconferencing, rather than (for example) audiographic conferencing, because of the latter's emphasis on graphics rather than visuals in general.

A measure of the success of a teleconferencing system is how well the interaction between participants simulates that available in a face-to-face meeting. We shall suggest in this paper that such interaction may even be enhanced by appropriately applied technology.

II. Objectives, Requirements, and Constraints

The task-domain objective for a visual teleconferencing system may be stated as follows. The system should allow participants in teleconferences to share visual materials as effectively as they can in conventional meetings.

This objective leads to the formulation of requirements for a conferencing system. We list the key requirements below.

1) Should allow any number of participants at any number of sites.
2) Should allow all participants to be peers. (As the number of participants increases, the opportunities to interact will decrease and social protocols will dictate the conduct of the meeting. Nevertheless, the technology should not impose an artificial hierarchy on participants.)
3) Should allow the sharing of all visual media used in meetings. Essential are paper and a blackboard/whiteboard facility. Natural images and computer graphics and screens are also required.
4) Should allow the interactive updating and annotation of all visual media, and the saving and printing of visuals at any location.
5) Should be possible both to organize and deliver formal, highly structured presentations, and to introduce visuals out of sequence and without warning from any site.

The above requirements are to be met subject to constraints. The following constraints arise out of the desires to make a system a widely applicable and as economical as possible.

1) Should be realizable on popular desktop computers.
2) Should allow arbitrary combinations of optional peripherals for media input and output.
3) Should be independent of the communication network.
4) Should be usable after a short demonstration. The availability of advanced features should not hinder novice users.

Paper approved by the Editor-in-Chief of the IEEE Communications Society. Manuscript received December 15, 1988; revised January 15, 1990.
J. Robinson is with the Department of Systems Design Engineering, University of Waterloo, Waterloo, Ont. N2L 3G1, Canada.
E. Rubinov, C. Toulson, B. Prasada, S. Sabri, and G. Vonderweidt are with Bell-Northern Research Ltd., Verdun, P.Q. H3E 1H6, Canada.
N. Goldberg is with the Armament Development Authority, 31021 Haifa, Israel.
IEEE Log Number 8101568.

In the following sections we discuss previous work addressed at visual teleconferencing, then describe a system designed to meet the above ojective and requirements.

III. Prior Work

Two strands of prior work have particular importance to the objective stated above. The first is conventional visual teleconferencing; the second is computer supported cooperative work (CSCW).

Visual teleconferencing has traditionally relied on special-purpose hardware terminals to provide for multipoint distribution of visual material. A detailed review of such systems, including telewriting, elecontronic blackboards, freeze-frame and document conferencing systems is given in [1]. In some cases, only computer text and graphics are supported (examples are the Gamma Telesketch and the Telewriter II PC [2]), while others allow the display of freeze-frame video [3], [4] or scanned documents [5], [6]. Increasingly, several media have been combined into a single system. Sometimes telewriting (freehand graphics) is combined with audio [7]. Other alternatives are facsimile plus telewriting [8] and freeze-frame video plus computer text and graphics [9]. Several truly multimedia systems have been reported in the literature [10]–[14]. These exploit the advances in graphic displays mentioned above, and in some cases combine motion video with the traditional media (INVITE-64 [11]). Most visual conferencing terminals allow pointing on, and annotation of media, though even in advanced systems, users can often only do this one at a time [14]. Some systems are intended for use over broadband networks ([12], [13]), while others are ISDN terminals. In all cases, they use special-purpose hardware, though the terminal is often intended to be used for personal computing and other communications tasks as well as conferencing.

In our own earlier work on visual teleconferencing, we developed a model of the processes involved in interactive visual communications [15]. This provided a framework within which to design our early experimental systems [16], and its influence is seen in various aspects of MICA. A major emphasis of the model was the assumption that "the participants in a visual conversation should have the same scene in their field of view at all times; thus new visual information generated by one participant should be communicated immediately to the other participants." This led to the definition of a shared visual space, based on Thompson's common information space [17], and retained in MICA, as we shall see below.

The second strand of prior research is computer supported cooperative work (CSCW), a field of increasing interest today (see [18] and [19] for two somewhat different perspectives on CSCW). It differs from traditional visual teleconferencing in two major ways. First, it deals with communication between general-purpose computer workstations that have standard hardware, but may differ from each other. Second, it emphasizes native computer media rather than paper, vugraphs, or photographs. In addition, CSCW, still in its infancy, is dealing with a variety of human behavioral issues about how computers may be used to support cooperating (or even noncooperating) groups.

Several commercially available personal computer programs may be termed CSCW applications. These include group writing tools [20] and application sharing [21]. The latter is of particular interest since it allows the simultaneous viewing of live (i.e., active) applications on multiple workstations. Thus a user of a word processor or spreadsheet (for example) may show work-in-progress to remote colleagues, and make modifications as they watch.

MICA builds on earlier work in visual teleconferencing in its support of several common visual media (multimedia), real-time simultaneous freehand and structured annotation (interactive), and multipoint working (conferencing) over long distances. From CSCW, it borrows the idea of using standard workstations rather than special-purpose hardware. By the same token, a variety of configurations must be supported, including options for document and video input. Thus, rather than a hardware-specific system, MICA is an hardware-flexible application. MICA also addresses native computer media, allowing the sharing of captured screens and windows, text files, and, with an appropriate hardware configuration, live applications.

The technical issues that arose in the development of MICA include the capture, display and compression of the supported media, the user interface design and the multipoint communications service (and protocol). In each of these areas we have new results to report. Most are dealt with in this paper. Providing a multipoint service that allows simultaneous interaction, consistency control and effective handling of the point-to-point case, all over long distances, is considered separately in our companion paper [22]. A further publication [23] details a study into the display of natural images on graphics displays as used in MICA.

IV. Overview of MICA

A. MICA Services and Features

MICA allows people at two or more sites to hold meetings at a distance and share visual material such as slides, documents and photographs. Visuals are displayed on a high or medium-resolution screen in an area called the shared space, which occupies most of the screen. Within the shared space, all connected sites display the same thing. Users at any site can introduce visuals, point and annotate on the shared space, and their actions are made visible at all sites. MICA allows free interaction, simultaneous annotation and control of the conference by any participant. Access to the control functions is via graphical menus displayed beside the shared space.

The inital implementation of MICA runs on an IBM PC/AT-compatible personal computer (with either an 80286 or an 80386 processor), with display, communications and input hardware. There is no special-purpose hardware; the system is intended as a workstation application rather than a stand-alone conference tool. The user can choose either a medium-resolution (640 × 480 pels, VGA standard) or a high-resolution (1024 × 1024) display, either a mouse or a tablet and stylus, either a LAN, ISDN, or X.25 network interface. A document scanner, printer and video camera may be added if required.

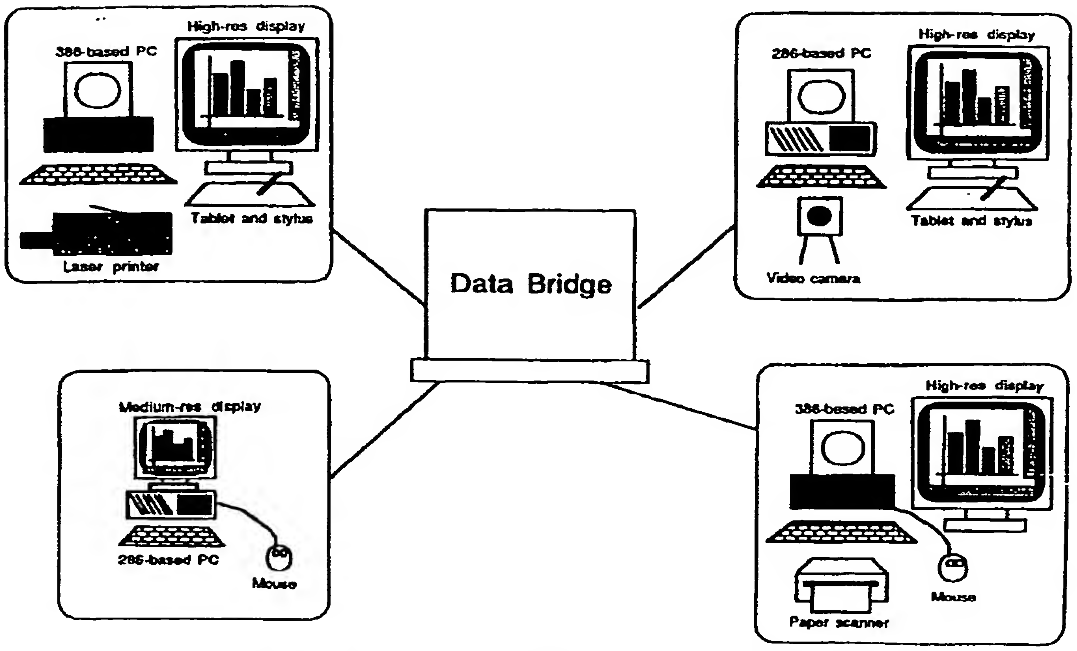

Fig. 1. Example MICA conference showing four sites connected to a databridge; different equipment configurations at each site, but the same view of the shared space on each screen.

Whatever hardware configuration a user has (and the alternatives above represent only the initial implementations), that system can receive and display all media, receive and display all annotations, communicate multipoint and interwork with all other configurations. Furthermore, all sites in a conference are peers; there are no master/slave relationships. Fig. 1 shows an example conference, illustrating the shared space concept and interworking between terminals of different hardware configuration.

The shared space can be used to display the current slide (a multimedia visual) or the conference monitor. The meeting is set up using the conference monitor, which also gives information about remote participants and the progress of the meeting. Most of the time the shared space is used for displaying slides; pre-stored slides for presentations, and documents and video images spontaneously input through the scanner and camera for unplanned contributions. Although the shared space looks identical at all participating sites, the menus may look different depending on the facilities selected by users. They include a tray of annotators (or pens) that users can pick up and draw with, and a lectern from which visuals are presented. Figs. 2 and 3 show the screen of the high-resolution version at different stages in a conference. Fig. 2 shows the conference monitor for a three-point meeting, with photographs of participants displayed (These appear in color, but are reproduced here in black and white). The buttons close to the bottom of the screen give access to special services. Fig. 3 shows a visual being displayed and annotated. A document has been scanned into the shared space and now a translucent highlighting pen is being used by one participant while the other writes with an opaque red pen. The lectern is also visible; this menu gives access to physical devices, together with three colored folders (red, green and blue) that represent stored sequenced slide presentations. The folders embody mechanisms (based on direct manipulation of miniature-image icons) for accessing, browsing and selecting slides for display on the shared space. Fig. 3 shows two open folders and one closed. Two miniature images are displayed for each open folder. These are content-generated icons that can be manipulated directly to flick through the folder or to present slides to the shared space. The miniatures seen in Fig. 3 represent, from left to right, visuals of the following media: GEM graphics, paper, a computer text file, and freeze-frame video, GEM graphics, captured PC screen, and paper.

B. MICA Software Architecture

MICA software is highly modular. This was a design priority so that addition of new media or new communications interfaces, changes in the screen resolution or the user interface, or any other modification, would have limited effect on the software. This in turn was a consequence of the design constraint that arbitrary combinations of optional peripherals should be allowed. Fig. 4 shows the MICA software modules in their five main groups: control, display interface, commu-

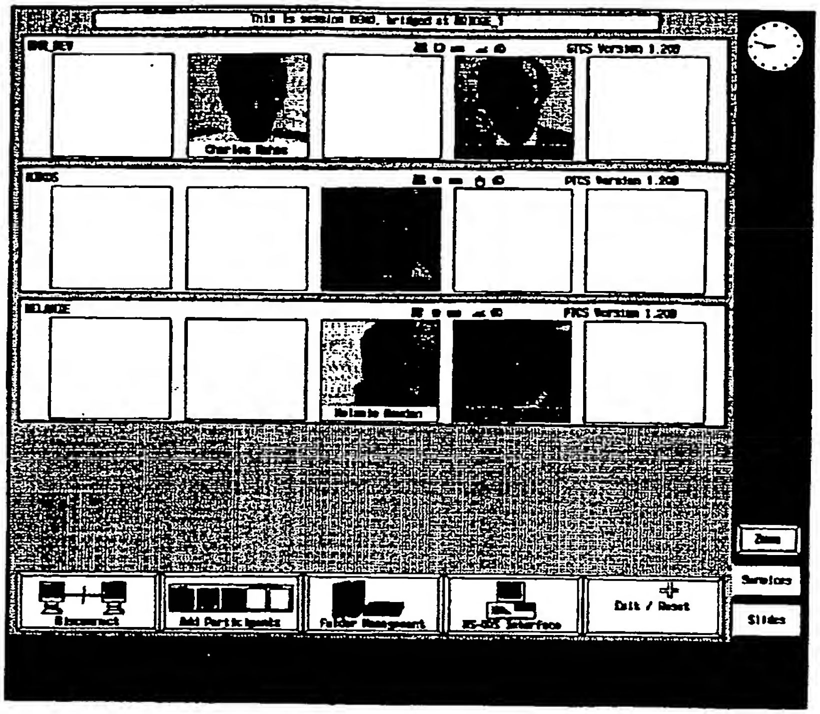

Fig. 2. Example of MICA services mode showing three sites connected. Site bars show participant pictures, with the control menus displayed across the bottom.

nications, multimedia processing and user interface. The lines between modules show the main channels of data transfer.

The control group's main function is to provide a non-preemptive task scheduler for MICA's six communicating processes. These are a user interface process, send and receive processes, a clock, and, when slides are being displayed, multimedia control and input processes. The use of multiple tasks allows us to provide concurrency between local events and remote events in a modular way. The scheduler also includes various functions that processes can call for synchronization purposes. The remainder of the first group of modules includes error reporting functions, initialization and debugging facilities.

The second group of modules provide the interface to the graphics display and the locater in a device independent way.

Communications software resides in a group of three modules; MCL, send and receive. MCL (multipoint communications layer) is the terminal implementation of a multipoint protocol intended for a wide range of applications (see below). It provides all the control for multipoint consistency, including functions for connecting to conferences and exchanging data. Send and receive are application-level communication modules. They provide the extra control, packetization, and message structure that is specific to MICA.

All user actions are interpreted by the user interface module which makes appropriate calls on other subsystems such as the conference monitor. The conference monitor includes data structures for maintaining the state of the meeting, both in terms of the sites that are connected (application level connection and disconnection), and the users at each site. It is accessed in a similar way by both the user interface and the receive modules.

The final group of modules is the multimedia processing group. The slide interface module is the controller for all major shared space events and ensures consistency both locally and globally. It uses media handlers to display the various slide types. These modules all involve medium-specific coding and display techniques. They interface to the communication modules via functions in slide interface that manage data transfer.

V. Multimedia Image Processing for Display and Compression

Table I shows the various media supported by MICA, how each is compressed and displayed with algorithms appropriate to its type.

Image compression is necessary to optimize end-to-end throughput and storage, and while standard coding is used where appropriate (for structured graphics and for documents), new algorithms have been developed for photographic images

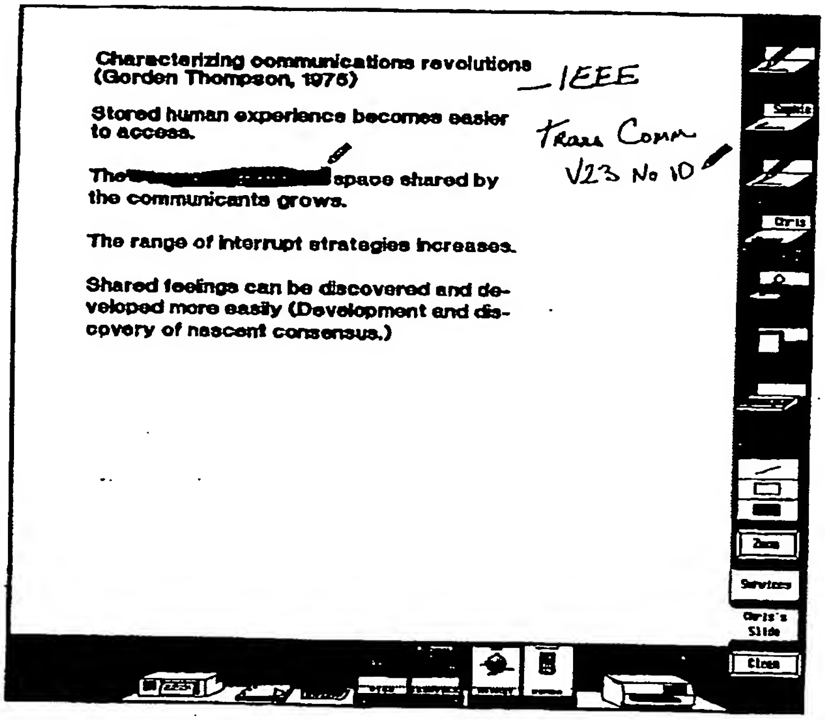

Fig. 3. Example of the MICA slides mode showing annotated shared space, folder menus with miniature slide icons, and the pen tray.

and PC screens. All media are coded according to their input resolution (for example, 300 dots/in for scanned documents), so the transmitted data are independent of the sender's or receiver's display capability. This is essential for compatibility between terminals with different display hardware, and also means that all media can be recovered at their original resolutions for printing.

Because the workstation's available display resolution and range of colors does not normally match the input image's resolution and range of colors, media-specific algorithms are used to render the image as faithfully as possible. MICA works with limited-palette graphics adapters that map the four or eight-bit resolution of the graphics memory into sixteen or 256 colors chosen from a large range of possible colors. The mapping is done by means of loadable lookup tables, and the set of colors used to display any particular image is known as the image palette. For documents, the palette includes several gray levels, used to compensate for the loss in spatial resolution. For photographic images, an image-adaptive palette selection algorithm is used.

A. Paper

Scanned documents are stored and transmitted by MICA in the T.6 format used in CCITT Group IV facsimile. This approach was taken for three reasons:

1) Group IV is a proven and efficient algorithm for transmission over reliable channels.
2) Stored documents may be generated and accessed by facsimile programs running on the workstation.
3) We have an interest in network-based services that may include document processing. For interworking with these, we chose to use a standard.

The document display algorithm uses gray levels to compensate for the reduction in spatial resolution from the 2550 × 3300 bilevel original, to the 928 or 464 pels available for display. The technique is akin to that for producing antialiased displays [24]. It involves dividing the input document into cells corresponding to the number of output (screen) pels, and counting the ratio of black to white points in each cell. This value is then quantized to an appropriate gray level for display. Efficient implementations of the algorithm make use of Bresenham's algorithm for arbitrary scaling between resolutions, and look-up tables for fast calculation. The term SARX (spatial-amplitude resolution exchange) has been coined for this process.

B. Freeze-frame Video

The image compression algorithm exploits the redundancy in the input picture to reduce the number of bits which are required to represent it from 24 b/pel to an average of 1.5

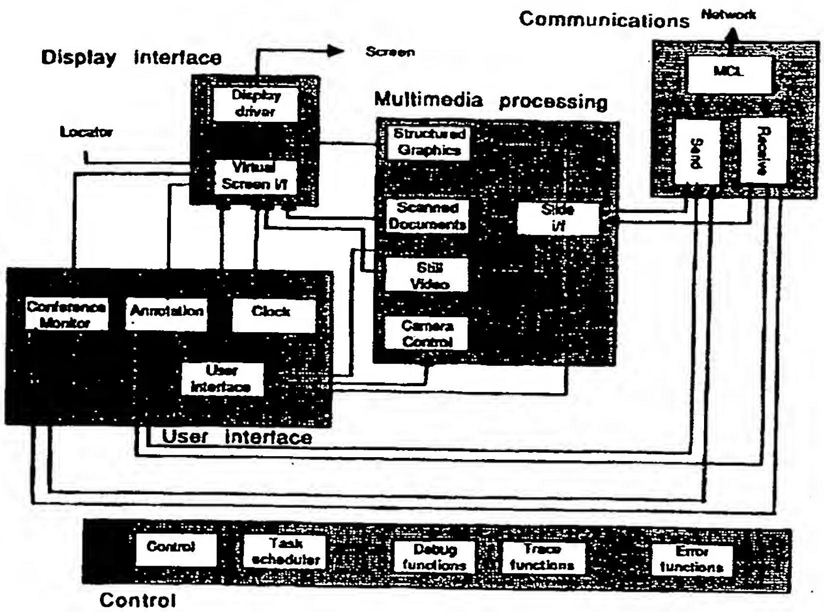

Fig. 4. Software architecture of MICA showing the five groups of modules and the major channels of data transfer.

TABLE I
THE MEDIA SUPPORTED BY MICA, THEIR COMPRESSION AND DISPLAY ALGORITHMS

Medium	Used for	Compressed by . . .	Displayed with . . .
Paper	Documents, Vugraphs	CCITT Group IV facsimile coding	Multi-grey-level algorithm
Freeze-Frame Video	Photographs, Objects, People pictures	Laplacian pyramid/predictive coding	Interative pallette selection, color quantization, dithering
GEM	Color graphics Graphs	Uncompressed GEM primitives	Resolution-independent graphics and fonts.
Captured PC Screens	Screens from application programs	Text screens not compressed. Multi-level run-length coding for graphics	Bit-mapped character fonts and raster graphics
ASCII Text	Unformatted PC text files	Not compressed: ASCII codes	Bit-mapped character fonts
Dynamic PC Screens	Live text-based PC applications	Conditional replenishment	Bit-mapped character fonts

b/pel, with slight degradation in the picture quality. The compression scheme is based on a combination of predictive and Laplacian Pyramid coding [25]. The motivation is to combine the fast implementation of predictive coding with the progressive transmission feature of the Laplacian Pyramid. All processing is done on the YIQ transformed color space, with subsampling in the I and Q components.

The algorithm is based on recursive generation of prediction error. In each step, a higher-resolution picture is predicted by interpolation of a lower-resolution picture. The lowest-resolution picture in our implementation is based on 8×16 blocks. The four pels at the corners of blocks are transmitted first and the remainder linearly interpolated at both the transmitter and receiver. The prediction errors between the interpolated block and the original are calculated at the transmitter, quantized and transmitted with variable word-length codes (runs of zero-error pels are also coded with run lengths).

Implementation of progressive transmission is straightforward; an entire low-resolution image is derived first by the linear interpolation from block-corner pels. The prediction errors, received next, serve to add detail. Having thus sent a monochrome image, the color components are transmitted in a similar way. In practice, progressive transmission has only been used in MICA systems communicating at 4.8 kb/s or below.

The display of a natural image on a graphics adapter requires the creation of an appropriate palette and efficient color selection from the palette during display. A survey of palette selection algorithms in the literature yielded three alternative approaches. In each case the first step is to generate the color histogram of the image. This is a 3-D matrix whose elements represent the RGB color space. Each cell in the histogram contains the number of occurrences of the corresponding color

in the picture. The first approach, the *Popularity* algorithm, simply selects the N most frequently occurring colors (for an N-entry palette) [26]. A modified version of the algorithm decreases the count of a selected color's neighbors to prevent the concentration of palette colors in one part of the color space [27]. The second approach is to arrange that each color in the palette will represent an equal number of pels in the image [26]. This is the *Median-cut* algorithm. The third approach starts by finding the M most frequently occurring colors (where M is many times greater than N, the number of entries in the palette). These represent a volume in color space. The algorithm, termed *MaxMin* [28], then seeks to find N uniformly distributed colors on the boundary of this populated region. Colors within the region can then be represented by combinations of the palette colors. Using combinations means that the image must either be displayed in higher resolution than it was input (for example using 2×2 cells for each input pel), or color quantization errors must be carried from one pel to the next (error diffusion). After a detailed comparison of these alternatives we developed a modified version of the MaxMin algorithm which involves recursive improvement to the selected palette after the initial choice of colors. This work is reported in detail in [23]. In the high-resolution version of MICA, a fast algorithm is used to generate 2×2 palette combinations for each displayed pel. In the medium-resolution version, where the usable palette in only 10 colors, dither and error diffusion are used to achieve a greater subjective color range.

C. Structured Graphics

For storage and transmission of structured graphics we elected to use GEM primitives [29]. Our initial implementation of the medium has been GemDraw based, and it was therefore natural to read and send unaltered GEM files. The graphics drawing routines used for display of structured graphics are all our own implementation. They include handling of outline fonts, so text is arbitrarily scalable. Decoding of other graphics standards is relatively simple; the process consists mainly of table lookup transformation. We have also implemented display of Macintosh PICT files. Since our initial workstation engine is a PC, these files have to be transferred prior to the conference. However, our current work includes implementation of MICA on a Mac II, and this medium will then have greater importance.

D. Text Files

Text files are transmitted uncompressed. MICA incorporates a feature to transfer arbitrary files, or complete directories. These files are also transmitted without compression. Text is displayed using standard computer fonts. The need for compatibility between systems while displaying lines of reasonable length means that the medium-resolution version uses an 8×16 font for text files. This is a solution that needs refining; we expect to use enhanced antialiased fonts in the future.

E. Captured PC Screens

Text screens are transmitted uncompressed. Graphics screens are compressed with a multilevel run-length coding scheme. Display of PC screens presents resolution-related problems, some of which have yet to be satisfactorily resolved. Our screen capture program works under all modes of common IBM display adaptors (Monochrome, Hercules, CGA, EGA, VGA). However, the aspect ratio differences between certain modes mean that the vertical dimension sometimes has to be scaled when displayed. A SARX solution (see above) to this would be ideal. The choice of palette for SARX requires investigation, since there are a large number of possible color boundaries on typical EGA displays. A related problem is that on the medium-resolution version of MICA, the screen is only as large as some of the captured screens. The shared space does not fill the screen, and the captured version must therefore be scaled downwards. One possiblility for this is to subsample as necessary and hope for the best; this approach leads to unpleasant distortions. Another possibility is to display captured screens in monochrome and use SARX to give smooth boundaries. We are still investigating options here.

F. Dynamic PC Screens

We have initially implemented only text-based application sharing on 80386-based machines (that is, the sender must have a 80386 processor, receivers need not). Compression is by conditional replenishment; only changed parts of the screen are sent as updates. Display of dynamic screens is similar to display of static screens. We have chosen to allow annotation on dynamic screens, although if a character position is updated, the annotation that occupied that space is erased.

VI. The User's View of the System

Figs. 2 and 3 show the user's view of MICA running on a high resolution display. They show a large shared space (80% of the screen area in the high-resolution version, slightly less in medium-resolution), with control menus along the bottom and the right hand side. The menus are not complex, and functions are localized. For example, when visuals are being displayed, presentation facilities are along the bottom menu and annotation facilities along the side. Some of the icons provide access to submenus. For example, selecting a scanner icon opens a menu that allows the user to specify whether documents should scanned for display in portrait, landscape or automatic format (automatic format selection relies on landscape documents being pre-marked in the top right-hand corner).

Little empirical data is available for deciding how a Computer Supported Cooperative Work interface should look. For example, the question of users' access to annotators is discussed by Schneiderman: "When several users collaborate over an electronic document, should they all have independent cursors and the freedom to update concurrently, or should there be a single cursor and some protocol for 'passing the chalk'? Intermediate strategies such as multiple cursors for pointing, but only one for making changes are possible. We could debate the merits of each approach and even implement them, but a scientific approach to evaluation can be a strong stimulus to sharper thinking" [30]. In the absence of empirical data on users' models, we have followed the rule of maximizing flexibility wherever possible. Thus our solution to

the annotators problem is to allow several pens that can be used for pointing or annotation simultaneously. If future behavioral studies show that this flexibility leads to user confusion or dissatisfaction, we will limit flexibility appropriately. With this proviso about the lack of empirical data, we now discuss three major principles of the user interface design: emulating the use of visual media in conventional meetings, devising meeting metaphors for the various MICA services, and the use of content-generated icons.

A. Sharing Visual Media in Meetings

One of MICA's goals was to allow distributed participants to discuss visual material as well or better than in conventional meetings. Taking the model of a semiformal meeting, we examined the facilities that are provided in a conventional conference room, considering their flexibility, ease of use and consistency. We then looked at how MICA, as a universal visual display device, could embody the positive features of conventional equipment. We were also concerned that it be usable by multiple participants at a single site. In this section we review this analysis and state our conclusions.

The principal equipment for sharing visuals in today's meetings is the overhead projector. It provides for display of hand-drawn visuals and printed vugraphs. Anything reproducible on a photocopier can be displayed, so long as the acetate copy is of approximately letter size. The presenter can point on the visual itself or on the screen, and can annotate the visual in color. The size of the image can be varied by moving the projector, and when used in small meeting rooms the image is bright and clear (high resolution).

Thirty-five millimeter slides are used for photographic images and for professional presentations. Expert visual design can make very attractive 35 mm graphics. They can be prepared on computer systems, and are therefore easy to alter until transferred to film. Slides cannot be annotated. Although the size of picture can be altered, and the resolution is always very high, the image usually has to be viewed in a dimmed room.

Blackboards, whiteboards and flipcharts allow on-the-spot sketching of an idea. They are also used, in education in particular, for their original purpose: the progressive introduction of preconceived material. Blackboards allow a teacher to introduce verbal and visual material together.

Computer output is an increasingly important medium. Visual material such as alphanumeric tables (spreadsheets), statistical graphics, text, etc., are used in meetings; either in hard copy format, or increasingly, displayed and manipulated directly on the computer. Large-screen projection systems may be used for direct display. The image is relatively dim, and cannot normally be annotated.

It is often important in meetings to show particular 3-D objects. A special case is when a participant is to be introduced to the audience, though there are many other examples, especially in working meetings, where objects are displayed. A limitation of many meetings is the mismatch in size between the object under discussion (from a miniature mechanical part to a building) and the number of participants who must view it.

For all the visual material mentioned above, MICA supports appropriate media. It is intended also to provide similar or increased facilities for interaction, superior consistency between the handling of different media, and enhanced control. A MICA user may be limited in respect of input hardware (for example, a scanner is required to input documents), but all users can display all media.

In comparison with an overhead projector, MICA is able to provide equal display functionality except for varying the display size. Images are rendered very faithfully (by means of SARX) and they can be annotated as required. Because multiple sites are able to modify the image simultaneously, a higher degree of interaction is possible than with a conventional projector.

Photographic images can be manipulated (annotated, etc.) more powerfully in MICA than in a slide projector. Although the resolution of the display is limited by the standard video input resolution, the images are adequate for most applications where photographic images are discussed. On the other hand, 35 mm graphics slides need to be sharp and clear. In MICA the appropriate medium for such information is structured graphics. If 35 mm graphics were input via a slide scanner, they would suffer a perceptible loss in quality when displayed. Structured graphics are guaranteed to produce the sharpest image possible, whatever the display resolution.

Computer screen and applications can be displayed within MICA. The degradation that would result from reproduction on conventional presentation media is avoided.

Display of arbitrary objects is possible in MICA within the constraints of the available input devices. With a medium-quality video camera and zoom lens, it is possible to display 3-D objects with resolution appropriate to the details under discussion.

When used by a single person at each site, MICA allows easy control and interaction. This is superior to conventional equipment, where participants usually have to leave their seats to present or annotate displayed visuals. MICA does have the same problem when multiple participants are located at a single site and must share the locater device (mouse or tablet and stylus). It may be appropriate to develop a hardware interface for multiple locater input. When many users are at a single site, MICA also suffers the handicap of its fixed display size. We have interfaced the high-resolution system to a large-screen projector, which gives acceptable viewing by a large group. However, the cost of such equipment is high, and the image is not as bright or crisp as on a CRT.

In conclusion, we see that MICA provides a good distributed counterpart to the equipment used in local meetings. It is superior to conventional equipment in that, so long as input devices are available, all media can be handled, displayed, annotated and stored in a consistent way.

B. Meeting Room Metaphors

A review of typical visual conferencing systems [1]–[13] shows that user interface designs usually include easy-to-use menus, with selections made by clicking. However, the value of metaphor [31] has not been exploited previously. Direct

TABLE II
FEATURES OF CONVENTIONAL MEETINGS AND THE CORRESPONDING METAPHORS IN MICA

Meeting Feature	Teleconference Correlate	Meeting Metaphor
Meeting room: "presence" and participants	Conference monitor (list of participants)	Virtual conference table, photos and nametags of participants
Projector (slide or overhead)	Shared space display of visuals	Miniatures of slides placed on Shared Space by dragging.
Presentation Visuals	Stored sequence of images	Slides in ring binders; pages turned by dragging
Writing on Visuals	Drawing mode	Shared pens (including highlighter, liquid paper and eraser)
Pointing on Visuals	Movable pointer	Pens are pointers
Flipchart/whiteboard	Blank screen for annotations	"Flipchart" viewed on the shared space
Interaction	Multiple data streams or baton passing	Pens may be used simultaneously. Control is by human protocols
—	Conference setup and communication control	"Services" viewed on the shared space

manipulation of familiar meeting room objects (such as ring binders, pens and nametags) is used in MICA to provide an elegant metaphor-based user interface. Table II shows how the visual aspects of a conventional meeting can be provided for electronically by a teleconference correlate, and the metaphors that MICA uses to control these features.

We have adopted the action of dragging as the most important mode of control, because it is both an intuitive and powerful metaphor for the basic manipulations to be done on MICA objects. For example miniatures can be dragged to other positions in the ring binders to allow browsing and arrangement, they can be dragged to the printer for hard copy, and they can be dragged to the shared space for presentation.

C. Content-Generated Icons

The user interface includes content-generated icons to represent objects that the user can manipulate. The iconic forms are generated directly from the data of the corresponding objects. So icons of meeting participants are photographs of the appropriate person, and image-slide icons are accurate miniatures of the visuals they represent. The size of an icon shows its class or function, but its visible pattern shows its particular content. Users rely on iconic descriptions of objects rather than their names, because the icons are unique and instantly recognizable. Note that the use of miniatures to represent larger visuals is only one way in which icons can be content-generated. In general, a content-generated icon is simply one in which the *content* (rather than the class or function) of the object is represented in iconic form.

Miniatures are generated by the appropriate media-handling parts of MICA. The use of multiple grays (SARX) for documents means that even at a resolution of 64 × 64 pels, they yield sufficiently faithful icons, so that visual cuing is very fast. We have not done formal tests on the effectiveness of miniatures in fast browsing, but the enthusiastic response to them in MICA leads us to believe that other applications could benefit from their use.

Consider a user interface with conventional iconic properties plus the ability to scale text and image data within windows. The size of characters would be variable continuously, with controls associated with the window for zooming-in and zooming-out. Resizing the window may result in some automatic scaling of the contents to make best user of the available area. At certain points on the continuum of possible window sizes it will be appropriate to reduce the window's functionality in accordance with its prominence on the screen, so that in the limit the window becomes an icon with visual (but not functional) similarity to the large window version. The user may leave a window in a particular state so that its iconic form gives cues to its meaning. Thus the system allows many windows simultaneously displaying data, possibly in several different ways (e.g., text, scanned document, graphics, photographic).

In a graphical interface where files are shown with content-generated icons, the operation of zooming the icon out into a window will start up the application for processing the file. This has implications for the way applications are written, at least for a multitasking environment. Applications should not include functions for opening and closing documents; rather opening a document should cause the creation of a new data area for a re-enterable application that may already be handling a number of other documents.

VII. TRIALS EXPERIENCE, IMPLICATIONS AND CONCLUSION

MICA has been demonstrated in several forums and has been the object of a number of user trials in tele-education and ISDN. User feedback has been overwhelmingly positive, with the chief criticism being the lack of hard copy. This has now been rectified. Other comments have related to partitioning and zooming of the Shared Space, and advanced user interface features (for more expert users). On hardware side, the issue of a larger display has been raised, along with multiple input locaters at a single MICA station.

As discussed earlier, MICA can be interfaced to a variety of communication networks. Specifically, it has been used over PC-net (a LAN), ISDN D (16 kb/s) and B (64 kb/s) channels, Datapac (X.25 packet data network), Datalink (circuit-switched data service), 56 kb/s leased lines and the

TABLE III
COMMUNICATIONS OPTIONS FOR MICA

Network	Data Rate	Time to Transmit Typical Document Image Between 80286-Based Terminals	Comments
PC-Net (LAN)	2 Mb/s	15 s (processor limited)	Limited reach
ISDN B-channel (HDLC, X.25)	64 kb/s	15 s (processor limited)	Future standard; not widely available
X.25 over a packet network	19.2 kb/s 9.6 kb/s	15 s (processor limited) 20 s (bandwidth limited)	Node delays were negligable in all tests to date
X.25 over circuit-switched network	9.6 kb/s 4.8 kb/s 2.4 kb/s	20 s (bandwidth limited) 40 s (bandwidth limited) 80 s (bandwidth limited)	Figures apply to dedicated data services and to transmission over conventional telephone lines. Noisy lines result in transmission time increases

PSTN. Table III shows the trade-offs in using different communication options. In particular, it shows that the time required to transmit a slide depends not just on the channel data rate, but also on the complexity of the image and the processing power of the PC. In most cases however, the choice of communication option will depend on the environment in which MICA is being installed.

In conclusion, we are also seeking to understand market requirements for the many kinds of computer-based teleconferencing. We are examining several vertical markets, of which tele-education is one, with a view to providing enhanced communication via MICA technology. The future success of multimedia conferencing products will depend on a good understanding, by developers, of the social, behavioral and economic aspects of the target user environment.

REFERENCES

[1] K. Watanabe, S.-I. Muakami, H. Ishikawa, and T. Kamae, "Audio and visually augmented teleconferencing," *Proc. IEEE*, vol. 73, pp. 656–670, Apr. 1985.
[2] *Telewriter II PC*, Optel Communications Inc., 1985.
[3] W. B. Pennebaker *et al.*, "A high-performance freeze frame video conferencing system," *Proc. GLOBECOM*, 1983, pp. 554–561.
[4] H. Yamaguchi, K. Yamada, K. Maya, and I. Iidaka "A graphics system for international teleconference," in *Proc. Int. Teleconf. Symp.*, Apr. 1984, pp. 221–226.
[5] T. Omachi, "High-resolution image system for teleconferencing," in *Proc. ICC*, 1982, pp. 2G3.1–2G3.5.
[6] H. Kotera and H. Tanigawa, "Design of a document oriented teleconferencing system," in *GLOBECOM, Conf. Rec., 1987*, pp. 713–717.
[7] T. Kishimoto and Y. Sato, "Simultaneous transmission of voice and handwriting signals: Sketchphone," *IEEE Trans. Commun.*, vol. COM-29, 1981.
[8] T. Kamae, "Sketchfax—A terminal having telewriting and facsimile capabilities," in *Proc. ICC*, vol. 2, pp. D3.4.1–D3.4.5, June 1983.
[9] *The Optel PC Conferencing System-Telewriter 3-PC*, Optel Communications Inc.
[10] T. Mikami, H. Takai, and N. Hayashi, "The Development of multimedia terminals in NTT," *IEEE Trans. Consumer Electron.*, vol. CE-31, pp. 323–331, Aug. 1985.
[11] C. N. Judice, and D. LeGall, "Telematic services and terminals: Are we ready?," *IEEE Commun. Mag.*, vol. 25, pp. 19–29, July 1987.
[12] M. Tasto, "Terminals for broad-band office communications," *IEEE J. Select. Areas Commun.*, vol. SAC-4, pp. 625–632, July 1986.
[13] M. Mera, K. Mitsuhashi, H. Shimizu, and S. Tsuruta, "Multimedia communication terminal for broadband network," in *GLOBECOM Conf. Rec.*, 1987, pp. 718–722.
[14] L. Aguilar, J. J. Garcia-Luna-Aceves, D. Moran E. J. Craighill, and R. Brungardt, "Architecture for a multimedia teleconferencing system," in *Proc. SIGCOMM '86*, Stowe, VT, Aug. 5–7 1986, pp. 126–136.
[15] M. D. Benjamin, R. D. Johnston, and B. Prasada, "A model for computer-mediated interactive visual communications," in *Proc. ICC '79*, June 1979, pp. 56.2.1–56.2.6.
[16] M. Ferguson, R. Johnston, J. Mastronardi, and B. Prasada, "IVC: Computer-mediated meeting-at-a-distance," *Telesis*, vol. 9, no. 1, pp. 8–13, 1982.
[17] G. B. Thompson, "An assessment methodology for evaluating communications innovations," *IEEE Trans. Commun.*, vol. COM-23, pp. 1045–1054, Oct. 1975.
[18] M. H. Olson, Ed., Technological support for work group collaboration, *1987 NYU Symp.*, May 21–22, 1991, Cent. Res. Inform. Syst. New York Univ.
[19] J. Nielsen, "Computer-Supported Cooperative Work," *ACM SIGCHI Bulletin*, vol. 19, no. 1, pp. 54–61, July 1987.
[20] L. S. Richman, "Software catches the team spirit," *Fortune*, pp. 125–136, June 8, 1987.
[21] "*In-Synch Desktop Teleconferencing*, *American Video Teleconferencing Corp.*
[22] G. Vonderweidt, J. Robinson, C. Toulson, J. Mastronardi, E. Rubinov, B. Prasada, and S. Sabri, "A multipoint communication service for interactive applications," *IEEE Trans. Commun.*, to be published.
[23] N. Goldberg, "Color image quantization for high resolution graphics display," submitted for publication in *Image and Vision Computing*.
[24] K. S. Booth, M. P. Bryden, W. B. Cowan, M. F. Morgan, and B. L. Plante, "On the parameters of human visual performance—an investigation of the benefits of antialiasing," In *CHI+GI 1987 Conf. Proc.*, Apr. 5–9, 1987, pp. 13–19.
[25] P. Burt, and E. Adelson, "the laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. COM-31, pp. 532–540, Apr. 1983.
[26] P. Heckbert, "Color quantization for frame buffer display," *Comput. Graph.*, vol. 16, pp. 297–307, 1982.
[27] G. Braudaway, "A procedure for optimum choice of a small number of colors from a large color palette for color imaging," *Electron. Imaging*, pp. 75–79, Nov. 1986.
[28] G. Houle and E. Dubois, "Quantization of color images for display on graphics terminals," in *GLOBECOM 1986 Conf. Rec.*, pp. 1138–1142.
[29] *Introduction to GEM Programming*, GEM Developers Kit, Digital Research, June 1986.
[30] B. Schneiderman, "Empirical studies in CSCW," in *CSCW Panel, CHI+GI 1987 Conf. Proc.*, Apr. 5–9, 1987, p. 228.
[31] D. C. Smith, C. Irby, R. Kimball, and B. Verplank, "Designing the start user interface," *Byte Magazine*, Apr. 1982, pp. 242–282.

John Robinson (M'87) received the B.Sc. degree from Durham University, England, in 1979, and the M.S. and Ph.D. degrees from the University of Essex, England, in 1982 and 1985, respectively.

He spent two years as a Development Engineer with Standard Telephones and Cables Ltd., Basildon, England, and three years as a Member of Scientific Staff, and then as a Manager, with Bell-Northern Research Ltd., Verdun, P.Q., Canada. He is now an Assistant Professor at the University of Waterloo, Waterloo, Ont., with research interests in visual communication and network services for computer-supported cooperative work.

Eliot Rabinov was born in Montreal, P.Q., Canada, in 1954. He received the honours degree in electrical engineering from McGill University, Montreal, in 1976, and the M.S.E.E. degree from Carnegie-Mellon University, Pittsburgh, PA, in 1979.

He joined Bell-Northern Research in 1977 as a member of Scientific Staff with the Digital Signal Processing group in Montreal. He contributed technically to various research activities in image communications, in particular to the development of experimental systems for multipoint interactive visual teleconferencing. Since 1985 he has managed research and development activities in the areas of multimedia teleconferencing, messaging and information services. He is currently manager of a department responsible for the definition and development of advanced network-based voice and image communication services.

Christopher Toulson was born in Yorkshire, England, on February 15, 1950. He received the honors degree in electrical engineering from McGill University, Montréal, P.Q., Canada, in 1972.

Prior to joining Bell-Northern Research, he spent several years working on advanced nuclear power station and flight simulators. He joined BNR Montréal in 1982 as Manager of the software systems group responsible for the network services department's computer facilities. In 1986 he was named Manager of the audiovisual bridging group, part of a team responsible for the design and implementation of an advanced teleconferencing research system. His research interests include interactive multimedia image handling and distributed multipoint applications.

Birendra Prasada (M'63–SM'76–F'89) received the B.Sc. and M.Sc. degrees in physics from Banaras Hindu University, the D.I.C. diploma in electrical engineering from Imperial College, London, and the Ph.D. degree in electrical communications from the University of London.

He is Director of Systems Research and Network Services at Bell-Northern Research and Professeur Invité at l'Institut National de la Recherche Scientifique (INRS)-Telecommunications, Université du Québec, Montréal, P.Q., Canada. His main research interests are in the areas of Multimedia communications and services, Digital Signal processing and human communication. His previous employment includes work at the Central Electronics Engineering Research Institute, Pilani, India, and the Defense Science Laboratory, Delhi, India, from 1961 to 1963. In 1963, he joined Bell Laboratories, Murray Hill, NJ as a Member of the Technical Staff, and in 1965 was appointed Assistant Professor in the Department of Electrical Engineering at the Massachusetts Institute of Technology, Cambridge. He joined the Indian Institute of Technology, Kanpur, India, in 1966, where he served as Professor and the Head of the Department of Electrical Engineering from 1968 to 1972, then as head of the Advanced Center for Electronic Systems from 1972 to 1973. From 1973 to 1976, he was a member of the Technical Staff of the Electronics and Computers Systems Research Laboratory of Bell Laboratories, Holmdel, NJ. From 1976 to 1985, he was Manager of Visual Communications Systems Research at Bell-Northern Research. He has also acted as a Consultant for industries in both India and the United States.

Shaker Sabri (M'76) was born in Alexandria, Egypt, on February 12, 1947. He received the B.Sc. and M.Sc. degrees in electrical engineering from Alexandria University, in 1968 and 1972, respectively, and the Ph.D. degree from the University of Ottawa, Ottawa, Ont., Canada, in 1977.

From 1968 to 1971 he worked with the Army signaling Corps on the design and implementation of communication networks. In 1971 he joined the Department of Electrical Engineering, Alexandria University, as a Lecturer, where he was involved in teaching and carrying out research in the digital signal processing area. From 1973 to 1976 he was with the Department of Electrical Engineering, University of Ottawa. In 1976 he joined BNR, Montreal, Canada where he has held several senior R&D management responsibilities. Presently he manages the Video Services and Systems Department in BNR.

Naftaly Goldberg was born in Poland in August 1950. He received the B.Sc. and M.Sc. degrees in electrical engineering from the Technion–Israel Institute of Technology, Haifa, in 1976 and 1980, respectively.

From 1976 to 1979 he was at Elscint Ltd., Haifa, where he was involved in development of signal processing for computer aided tomography. Since 1979 he has been with the Armament Development Authority, Haifa, where he has been responsible for research and development work in signal and image processing. From 1985 to 1987 he was on sabbatical leave at Bell-Northern Research, Montréal, P.Q., Canada, where he was involved in research and development of image processing for a teleconference system.

Guy Vonderweidt received the engineering degree (majoring in computer science) with honors from the Institut National des Sciences Appliquées, Lyons, France in 1977.

For the next three years, he worked at Informatech France–Québec, Montréal, P.Q., Canada, specializing in operating systems and large databases for information retrieval. He then became partner in a small company, designing microcomputer hardware and software for dental office management. In 1981, he joined Bell-Northern Research, Montreal, as a Member of Scientific Staff in the Interactive Visual Communication Department. He developed numerous data communication systems and designed algorithms for multipoint visual communication. In 1988, he became manager of multimedia research. In 1990, and part of 1991, he managed a telephony applications group in Northern Telecom Meridian R & D Center, Marne La Vallée, France. He is currently back with BNR, Montreal.

data handling; data compression; image processing; multimedia interactive conferencing application; personal workstations; multipoint visual teleconferencing; MICA; documents; photographs; computer screens; real-time interaction; pointing; annotation; meeting services; multimedia teleconferencing; computer-supported cooperative work; user interface; computerised picture processing; graphical user interfaces; interactive systems; microcomputer applications; multimedia systems; telecommunications computing; teleconferencing

---